Safe Equilibrium

Anonymous Author(s)
Affiliation
Address
email

Abstract

The standard game-theoretic solution concept, Nash equilibrium, assumes that all players behave rationally. If we follow a Nash equilibrium and opponents are irrational (or follow strategies from a different Nash equilibrium), then we may obtain an extremely low payoff. On the other hand, a maximin strategy assumes that all opposing agents are playing to minimize our payoff (even if it is not in their best interest), and ensures the maximal possible worst-case payoff, but results in exceedingly conservative play. We propose a new solution concept called safe equilibrium that models opponents as behaving rationally with a specified probability and behaving potentially arbitrarily with the remaining probability. We prove that a safe equilibrium exists in all strategic-form games (for all possible values of the rationality parameters), and prove that its computation is PPAD-hard.

1 Introduction

In designing a strategy for a multiagent interaction an agent must balance between the assumption that opponents are behaving rationally with the risks that may occur if opponents behave irrationally. Most classic game-theoretic solution concepts, such as Nash equilibrium (NE), assume that all players are behaving rationally (and that this fact is common knowledge). On the other hand, a maximin strategy plays a strategy that has the largest worst-case guaranteed expected payoff; this limits the potential downside against a worst-case and potentially irrational opponent, but can also cause us to achieve significantly lower payoff against rational opponents. In two-player zero-sum games, Nash equilibrium and maximin strategies are equivalent (by the minimax theorem), and these two goals are completely aligned. But in non-zero-sum games and games with more than two players, this is not the case. In these games we can potentially obtain arbitrarily low payoff by following a Nash equilibrium strategy, but if we follow a maximin strategy will likely be playing far too conservatively. While the assumption that opponents are exhibiting a degree of rationality, as well as the desire to limit worst-case performance in the case of irrational opponents, are both desirable, neither the Nash equilibrium nor maximin solution concept is definitively compelling on its own.

We propose a new solution concept that balances between these two extremes. In a two-player general-sum game, we define an ϵ*-safe equilibrium* (ϵ-SE) as a strategy profile where each player i is playing a strategy that minimizes performance of the opponent with probability ϵ_i, and is playing a best response to the opponent's strategy with probability $1-\epsilon_i$, where $\epsilon = (\epsilon_1, \epsilon_2)$. As a special case, if we are interested in constructing a strategy for player 1, we can set $\epsilon_1 = 0$, assuming irrationality just for player 2. We can generalize this to an n-player game by assuming that all players $i \neq 1$ are playing a strategy that minimizes player 1's expected payoff with probability ϵ_i, and are playing a best response to all other players' strategies with probability $1-\epsilon_i$, while player 1 plays a best response to all other players' strategies. This concept balances explicitly between the assumption of players' rationality and the desire to ensure safety in the worst case through the ϵ_i parameters.

Several other game-theoretic solution concepts have been previously proposed to account for degrees of opponents' rationality. The most prominent is *trembling-hand perfect equilibrium* (THPE), which

Submitted to 37th Conference on Neural Information Processing Systems (NeurIPS 2023). Do not distribute.

is a refinement of Nash equilibrium that is robust to the possibility that players "tremble" and play each pure strategy with arbitrarily small probability [4]. The concept of ϵ-safe equilibrium differs from THPE in several key ways. First, it allows a player to specify an arbitrary belief on the probability that each other player is irrational, rather than assume that it is an extremely small value. In domains like national security or driving we risk losing lives in the event that we fail to properly account for opponents' irrationality, and may elect to use larger values for ϵ_i than in situations where safety is less of a concern. In an ϵ-SE a player can specify the values for ϵ_i based on prior beliefs about the opponent or any relevant domain-specific knowledge, and is still free to use values that are extremely close to 0 as in THPE. Furthermore, a THPE is a refinement of NE, while ϵ-SE and NE are incomparable (an ϵ-SE may not be an NE and vice versa). Another related concept is that of a *safe strategy* and *ϵ-safe strategy* [3]. A strategy for a player in a two-player zero-sum game is called safe if it guarantees an expected payoff of at least v^*—the value of the game to the player—in the worse case. Note that this also coincides with the set of minimax, maximin, and Nash equilibrium strategies. A strategy is ϵ-safe if it obtains a worst-case expected payoff of at least $v^* - \epsilon$. The concepts of safe and ϵ-safe strategies are defined just for two-player zero-sum games, while safe and ϵ-safe equilibrium also apply to non-zero-sum and multiplayer games.

We note that a belief of opponents' "irrationality" does not necessarily indicate that we believe them to be "stupid" or "crazy." It may simply correspond to a belief that the opponent may have a different model of the game than we do. For example, our analysis may indicate that a successful attack on a location would result in a certain payoff for the opponent, while their analysis indicates a different payoff. In addition to potentially constructing different assessments of their own or other players' payoffs, opponents may also be "irrational" because they are using an algorithm for computing a Nash equilibrium that is only able to yield an approximation, or just a different Nash equilibrium from what other players have calculated (in fact, these cases do not actually seem to be irrational at all, since computing a Nash equilibrium is computationally challenging and many games have multiple Nash equilibria). If any of these situations arise, then simply following an arbitrary Nash equilibrium strategy runs a risk of an extremely low payoff, and there is potential for significant benefit by ensuring a degree of safety.

An alternative approach for modeling potentially irrational opponents is to incorporate an *opponent modeling algorithm*. An approach called a restricted Nash response was developed for two-player zero-sum games where the opponent is restricted to play a fixed strategy σ_{fix} determined by an opponent model with probability p and plays a best response to us with probability $1 - p$ while we best respond to the opponent (it is shown that this approach is equivalent to playing an ϵ-safe best response to σ_{fix} (a best response to σ_{fix} out of strategies that are ϵ-safe) for some ϵ) [2]. It was shown that for certain values of p this approach can result in a significant reduction in the level of exploitability of our own strategy while only a slight reduction in our degree of exploitation of the opponent's strategy. It has also been shown that approaches that compute an ϵ-safe best response to a model of the opponent's strategy for dynamically changing values of ϵ in repeated two-player zero-sum games can guarantee safety [1]. An ϵ-safe equilibrium strategy can be used in non-zero-sum and multiplayer games where models are available for the opponents' strategies by assuming each opponent i follows their opponent model with probability ϵ_i instead of playing a worst-case strategy for us, while also playing a best response with probability $1 - \epsilon_i$. Thus, in the event that an opponent model is available we can view safe equilibrium as a generalization of restricted Nash response to achieve robust opponent exploitation in the settings of non-zero-sum and multiplayer games.

2 Safe Equilibrium

A *strategic-form game* consists of a finite set of players $N = \{1, \ldots, n\}$, a finite set of pure strategies S_i for each player $i \in N$, and a real-valued utility for each player for each strategy vector (aka *strategy profile*), $u_i : \times_i S_i \rightarrow \mathbb{R}$. A *mixed strategy* σ_i for player i is a probability distribution over pure strategies, where $\sigma_i(s_{i'})$ is the probability that player i plays pure strategy $s_{i'} \in S_i$ under σ_i. Let Σ_i denote the full set of mixed strategies for player i. A strategy profile $\sigma^* = (\sigma^*_1, \ldots, \sigma^*_n)$ is a *Nash equilibrium* if $u_i(\sigma^*_i, \sigma^*_{-i}) \geq u_i(\sigma_i, \sigma^*_{-i})$ for all $\sigma_i \in \Sigma_i$ for all $i \in N$, where $\sigma^*_{-i} \in \Sigma_{-i}$ denotes the vector of the components of strategy σ^* for all players excluding player i. Here u_i denotes the expected utility for player i, and Σ_{-i} denotes the set of strategy profiles for all players excluding player i. A mixed strategy σ^*_i for player i is a *maximin strategy* if $\sigma^*_i \in \arg\max_{\sigma_i \in \Sigma_i} \min_{\sigma_{-i} \in \Sigma_{-i}} u_i(\sigma_i, \sigma_{-i})$.

Definition 1. *Let G be a two-player strategic-form game. Let $\epsilon = (\epsilon_1, \epsilon_2)$, where $\epsilon_i \in [0,1]$ for $i = 1, 2$. A strategy profile σ^* is an* ϵ-safe equilibrium *if there exist mixed strategies $\tau_i^*, \rho_i^* \in \Sigma_i$ where $\sigma_i^* = \epsilon_i \tau_i^* + (1-\epsilon_i)\rho_i^*$ for $i = 1, 2$ such that $\rho_i^* \in \arg\max_{\sigma_i \in \Sigma_i} u_i(\sigma_i, \sigma_{-i}^*)$, $\tau_i^* \in \arg\min_{\sigma_i \in \Sigma_i} u_{-i}(\sigma_{-i}^*, \sigma_i)$.*

In practice player i would likely want to set $\epsilon_i = 0$ and $\epsilon_j > 0$ for $j \neq i$ when determining their own strategy, though Definition 1 allows an arbitrary value of $\epsilon_i \in [0,1]$ as well. It may make sense for player i to set $\epsilon_i > 0$ if they believe both that the opponent is irrational with some probability ϵ_{-i}, and if they also believe that the opponent believes that player i is irrational with some probability ϵ_i.

Theorem 1. *Let $G = (N, (S_i)_{i \in N}, (u_i)_{i \in N})$ be a two-player strategic-form game, and let $\epsilon = (\epsilon_1, \epsilon_2)$, where $\epsilon_1, \epsilon_2 \in [0,1]$. Then G contains an ϵ-safe equilibrium.*

Proof. Define $G' = (N', (S'_i)_{i \in N}, (u'_i)_{i \in N})$ to be the following game. $N' = \{1,2,3,4\}$, $S'_1 = S'_2 = S_1$, $S'_3 = S'_4 = S_2$. For $s'_i \in S'_i$, define u'_i as follows for $i \in N$:

$$u'_1(s'_1, s'_2, s'_3, s'_4) = -\epsilon_2 u_2(s'_1, s'_3) - (1-\epsilon_2) u_2(s'_1, s'_4)$$
$$u'_2(s'_1, s'_2, s'_3, s'_4) = \epsilon_2 u_1(s'_2, s'_3) + (1-\epsilon_2) u_1(s'_2, s'_4)$$
$$u'_3(s'_1, s'_2, s'_3, s'_4) = -\epsilon_1 u_1(s'_1, s'_3) - (1-\epsilon_1) u_1(s'_2, s'_3)$$
$$u'_4(s'_1, s'_2, s'_3, s'_4) = \epsilon_1 u_2(s'_1, s'_4) + (1-\epsilon_1) u_2(s'_2, s'_4)$$

Player 1's strategy corresponds to τ_1^*, player 2's strategy corresponds to ρ_1^*, player 3's strategy corresponds to τ_2^*, and player 4's strategy corresponds to ρ_2^*. By Nash's existence theorem, the game G' has a Nash equilibrium, which corresponds to an ϵ-safe equilibrium of G. □

Theorem 2. *The problem of computing an ϵ-safe equilibrium in two-player games is PPAD-hard.*

Proof. Let $G = (N, (S_i)_{i \in N}, (u_i)_{i \in N})$ be a two-player strategic-form game. Suppose that k is the smallest possible payoff for any player in G, and let $k' = k - 1$. Define the game $G' = (N', (S'_i)_{i \in N}, (u'_i)_{i \in N})$ as follows. $N' = \{1,2\}$, $S'_1 = S_1 \cup t$, $S'_2 = S_2 \cup t$. For $s'_i \in S'_i$, define u'_i as follows for $i \in N$:

$$u'_i(s'_1, s'_2) = u_i(s'_1, s'_2) \text{ for } s_1 \in S_1, s_2 \in S_2.$$
$$u'_i(t, s'_2) = k' \text{ for } s'_2 \in S_2.$$
$$u'_i(s'_1, t) = k' \text{ for } s'_1 \in S_1.$$
$$u'_i(t, t) = k'.$$

Suppose we can efficiently compute an ϵ-safe equilibrium of G', denoted by $\sigma^{G'}$. Then we have $\sigma_i^{G'} = \epsilon_i \tau_i^* + (1-\epsilon_i)\rho_i^*$ for $i = 1, 2$, where $\rho_i^* \in \arg\max_{\sigma'_i \in \Sigma'_i} u_i(\sigma'_i, \sigma_{-i}^{G'})$, $\tau_i^* \in \arg\min_{\sigma'_i \in \Sigma'_i} u_{-i}(\sigma_{-i}^{G'}, \sigma'_i)$. I claim that ρ^* is a Nash equilibrium of G. First note that ρ_i^* must put probability 0 on t for all players, since t is strictly dominated. So it is a valid strategy profile of G. Also note that τ_i^* must put probability 1 on t for all i. Suppose that player i can improve performance in G by deviating to η_i. Then

$$u_i(\eta_i, \rho_{-i}^*) > u_i(\rho_i^*, \rho_{-i}^*)$$
$$(1-\epsilon_i) u_i(\eta_i, \rho_{-i}^*) + \epsilon_i k' > (1-\epsilon_i) u_i(\rho_i^*, \rho_{-i}^*) + \epsilon_i k'$$
$$(1-\epsilon_i) u_i(\eta_i, \rho_{-i}^*) + \epsilon_i u_i(\eta_i, t) > (1-\epsilon_i) u_i(\rho_i^*, \rho_{-i}^*) + \epsilon_i u_i(\eta_i, t)$$
$$(1-\epsilon_i) u_i(\eta_i, \rho_{-i}^*) + \epsilon_i u_i(\eta_i, t) > (1-\epsilon_i) u_i(\rho_i^*, \rho_{-i}^*) + \epsilon_i u_i(\rho_i^*, t)$$
$$(1-\epsilon_i) u_i(\eta_i, \rho_{-i}^*) + \epsilon_i u_i(\eta_i, \tau_{-i}^*) > (1-\epsilon_i) u_i(\rho_i^*, \rho_{-i}^*) + \epsilon_i u_i(\rho_i^*, \tau_{-i}^*)$$
$$u_i(\eta_i, \sigma_{-i}^{G'}) > u_i(\rho_i^*, \sigma_{-i}^{G'}).$$

This contradicts the fact that $\rho_i^* \in \arg\max_{\sigma'_i \in \Sigma'_i} u_i(\sigma'_i, \sigma_{-i}^{G'})$. So we have a contradiction, and conclude that no player can improve performance in G by deviating from ρ^*. So ρ^* is a Nash equilibrium of G. Since the problem of computing a Nash equilibrium is PPAD-hard and we have reduced it to the problem of computing an ϵ-safe equilibrium, this shows that the problem of computing an ϵ-safe equilibrium is PPAD-hard. □

$$\begin{bmatrix} (0,0) & (-1,+1) \\ (+1,-1) & (-10,-10) \end{bmatrix}$$

Figure 1: Payoff matrix for game of Chicken.

$$\begin{bmatrix} (4,-3) & (-1,1) & (-7,2) \\ (-5,5) & (2,-1) & (-1,4) \\ (-9,1) & (-1,8) & (9,-4) \end{bmatrix}$$

Figure 2: Security game payoff matrix.

For $n > 2$ players, we designate one of the players as being a special player, say player 1. We can view player 1 as representing "ourselves" as a decision-making agent, and the other players as unpredictable opponents. Player 1 then best responds to the strategy profile of all other players, while each opposing player i mixes between playing a strategy that minimizes player 1's payoff and a strategy that maximizes player i's payoff in response to the strategy profile of the other players.

Definition 2. *Let G be an n-player strategic-form game. Let $\epsilon = (\epsilon_2, \dots, \epsilon_n)$, where $\epsilon_i \in [0,1]$. A strategy profile σ^* is an ϵ-*safe equilibrium *if there exists a mixed strategy σ_1^* for player 1 and mixed strategies $\tau_i^*, \rho_i^* \in \Sigma_i$ where $\sigma_i^* = \epsilon_i \tau_i^* + (1-\epsilon_i)\rho_i^*$ for $i = 2, \dots, n$ such that $\rho_i^* \in \arg\max_{\sigma_i \in \Sigma_i} u_i(\sigma_i, \sigma_{-i}^*)$, $\tau_i^* \in \arg\min_{\sigma_i \in \Sigma_i} u_1(\sigma_1^*, \sigma')$, $\sigma_1^* \in \arg\max_{\sigma_1 \in \Sigma_1} u_1(\sigma_1, \sigma_{-1}^*)$, where σ' is the strategy profile for players 2–n where player i plays σ_i and the other players $j \neq i$ play σ_j^*.*

The proof of Theorem 1 extends naturally to $n > 2$ players as well by creating a $2(n-1)+1 = 2n-1$ player game with 2 new players corresponding to each original player for $i > 1$, plus player 1.

Theorem 3. *For all ϵ, every n-player strategic-form game contains an ϵ-safe equilibrium.*

Theorem 4. *The problem of computing an ϵ-safe equilibrium in n-player games is PPAD-hard.*

As an example, consider the classic game of Chicken, with payoffs given by Figure 1. The first action for each player corresponds to the "swerve" action, while the second corresponds to the "straight" action. The unique mixed-strategy Nash equilibrium σ^{NE} in the Chicken game is for each player to swerve with probability 0.9 (there are also two pure-strategy equilibria where one player swerves and the other player doesn't), and the unique maximin strategy σ^M is to swerve with probability 1. If we set $\epsilon_1 = 0$, then it turns out that σ_1^{NE} is an ϵ-safe equilibrium strategy for player 1 for $0 \leq \epsilon_2 \leq 0.1$, and σ_1^M is an ϵ-safe equilibrium strategy for player 1 for $0.1 \leq \epsilon_2 \leq 1$. It is not necessary that an ϵ-safe equilibrium strategy always corresponds to a Nash equilibrium or maximin strategy. For example, with $\epsilon_1 = 0.05$ and $\epsilon_2 = 0.15$, an ϵ-safe equilibrium strategy profile is for player 1 to swerve with probability 0.95 and player 2 to swerve with probability 0.

As another example, consider the security game depicted in Figure 2, where the row player selects one of three targets to defend while the column player selects a target to attack. A Nash equilibrium for player 1 (row player) σ_1^{NE} is to defend the targets with probabilities $(0.3136, 0.4661, 0.2203)$, and a maximin strategy σ_1^M is to defend the targets with probabilities $(0.6144, 0.0131, 0.3725)$. Again using $\epsilon_1 = 0$, for $\epsilon_2 \in [0, 0.314]$ it turns out that σ_1^{NE} is an ϵ-safe equilibrium strategy for player 1, and for $\epsilon_2 \in [0.569, 1]$ σ_1^M is an ϵ-safe equilibrium strategy for player 1. But for the region $\epsilon_2 \in [0.314, 0.569]$ it turns out that the strategy $(0.4437, 0.3666, 0.1897)$ is an ϵ-safe equilibrium strategy for player 1, which is neither a Nash equilibrium strategy nor a maximin strategy.

3 Conclusion

While Nash equilibrium has emerged as the central game-theoretic solution concept, its assumption that all players behave rationally may be too strict when modeling real human decision makers. As game theory is being increasingly applied to high-stakes situations, such as self-driving cars and national security, it is essential that strategies are able to accommodate the possibility of opponents' irrationality, which may be unpredictable. At the other end of the spectrum, a maximin strategy assumes that all opponents are trying to minimize our payoff, resulting in exceedingly conservative play with low payoffs. Safe equilibrium effectively bridges the gap between these two extremes, enabling us to construct strategies that are robust to arbitrary degrees of opponents' irrationality.

References

[1] Sam Ganzfried and Tuomas Sandholm. Safe opponent exploitation. *ACM Transactions on Economics and Computation (TEAC)*, 3(8):1–28, 2015.

[2] Michael Johanson, Martin Zinkevich, and Michael Bowling. Computing robust counter-strategies. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NIPS)*, pages 1128–1135, 2007.

[3] Peter McCracken and Michael Bowling. Safe strategies for agent modelling in games. In *AAAI Fall Symposium on Artificial Multi-agent Learning*, October 2004.

[4] Reinhard Selten. Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, 4:25–55, 1975.